



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM



FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

TIPO DE EVALUACIÓN: EXAMEN DIAGNÓSTICO.

MATERIA/UNIDAD DE APRENDIZAJE: Minería de datos.

I. Lee el caso de ejemplo y responde las preguntas correspondientes:

La Coca Cola Company tiene 366,000 accionistas y efectúa una encuesta por medio de la selección aleatoria de 30 accionistas de cada uno de los 50 estados de Estados Unidos. Se registra el número de acciones de cada accionista de la muestra.

- a) ¿Los valores obtenidos son discretos o continuos?

Discretos

- b) Identifique el nivel de medición (nominal, ordinal, de intervalo, de razón) de los datos muestrales.

Nominal

- c) ¿Qué tipo de muestreo (aleatorio, sistemático, de conveniencia, estratificado, por racimos) se utiliza?

Estratificado

- d) Si se calcula el número promedio (la media) de acciones. ¿El resultado es un estadístico o un parámetro?

Estadístico

- e) Si usted fuera ejecutivo en jefe de la Coca Cola Company, ¿qué característica del conjunto de datos consideraría que es extremadamente importante?

Que los datos dejan ver cómo se comportan las acciones por regiones, debido a como se obtuvo la información.

- f) ¿Qué es lo que está incorrecto al evaluar la opinión del accionista enviado un cuestionario por correo que los accionistas podrían llenar y regresar por el mismo medio?

Tal vez se puede perder la información de las respuestas de los encuestados.

II. Describe con tus palabras para que se utiliza el análisis uni-variante de datos:

Puede ser utilizado para recabar información muy específica de forma práctica, y así poder encontrar sus características de una forma más eficaz.

III. Describe qué es la media, la mediana y la desviación estándar:

La media es el valor que se obtiene al sumar todos los valores de un conjunto de datos y dividir dicha suma entre el número de observaciones, obteniendo así el valor promedio de los datos.

La mediana es el valor que se encuentra en la posición intermedia de todo el grupo de datos.



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM



FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

La desviación estándar es el valor que nos permite medir que tan dispersos están los datos con respecto a la media.

IV. Con tus palabras explica qué es minería de datos y en donde puede ser aplicada:

La minería de datos es un proceso que nos permite estudiar más de fondo elementos como bases de datos, para poder encontrar las características de los distintos grupos de datos, además de detectar anomalías.

V. Que lenguajes de programación dominas o en que porcentaje lo manejas:

Pues he usado visual basic, C++ y R. Siento que los primeros dos ahorita mismo estoy en rango del 55% al 70%, mientras que R apenas estoy aprendiendo

VI. Sabes manejar notebooks digitales como Jupyter, Spyder, Google Collab, etc. (Si la respuesta es otro, anota cuál):

Nunca he usado algún notebook digital.

VII. ¿Trabajas o has trabajado alguna vez en alguna empresa que maneje técnicas de ciencia de datos?

No he trabajado.

VIII. Has utilizado alguna herramienta de aprendizaje de máquina. Si la respuesta es afirmativa, anota cuál:

No.

IX. Relaciona los conceptos con sus gráficas correspondientes:

- a) Se utilizan para evaluar visualmente cómo se distribuyen los puntos de datos con respecto a su frecuencia.
- b) Forma estandarizada de mostrar la distribución de datos basada en un resumen de cinco números ("mínimo", primer cuartil (Q1), mediana, tercer cuartil (Q3) y "máximo"). Puede informarle si sus datos son simétricos, qué tan estrechamente están agrupados y si sus datos están sesgados.
- c) Método para visualizar la distribución de datos numéricos de diferentes variables. Es similar al diagrama de caja pero con un diagrama rotado en cada lado, que brinda más información sobre la estimación de densidad en el eje y.
- d) Visualización de datos bidimensionales que utiliza puntos para representar los valores obtenidos para dos variables diferentes: una trazada a lo largo del eje "xy" otra trazada a lo largo del eje y.



UANL

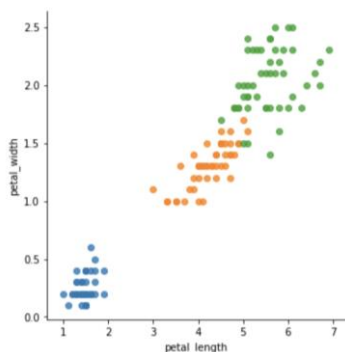
UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM

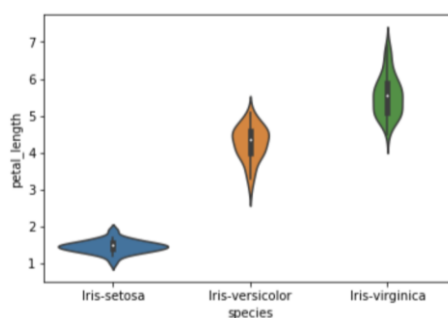
FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS



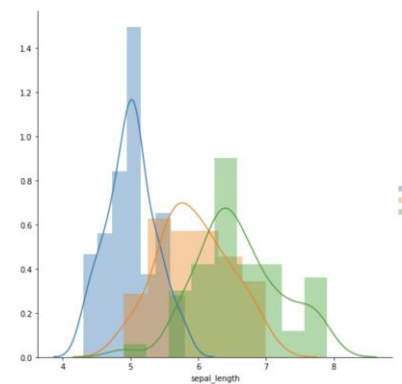
- e) Produce una matriz de relaciones entre cada variable en sus datos para un examen instantáneo de nuestros datos. También puede ser un gran punto de partida para determinar los tipos de análisis de regresión a utilizar.



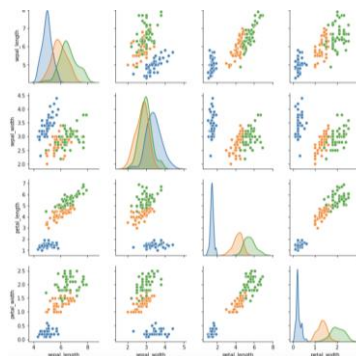
D



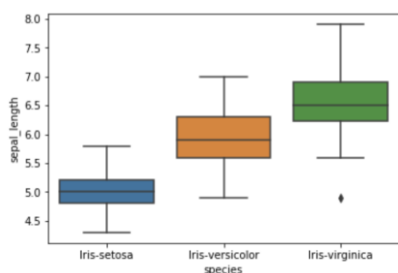
C



A



E



B



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM



FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

- X. Tienes problemas de compartir pantalla, o hablar en micrófono en el Teams: (Menciona cuales son los problemas en caso de tenerlos)

No tengo problemas.

- XI. Tienes problemas de abrir o utilizar Nexus: (Menciona cuales si tienes alguno)

Si puedo abrir Nexus, pero como no le he usado desde la prepa no se si funcione bien realmente.

- XII. Que sistema operativo tienes:

Windows

- XIII. De las siguientes herramientas menciona si tienes conocimiento de alguna y en que porcentaje.

- Github. No tengo conocimiento
- Jupyter Notebook No tengo conocimiento
- Google Collab No tengo conocimiento
- Python 3 No tengo conocimiento
- Terminal de tu computadora. No tengo conocimiento